BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

COMMONWEALTH LANGUAGE SERVICES, LTD.
Alison Sondhaus Carroll, President

6044 Little Falls Rd., Arlington, Virginia 22207
tel: (703) 534-7649 fax: (703) 534-4191
email: alisonscarroll@juno.com

MEMORANDUM

To:	Lara Kelley, esq./Brian Fish
From:	Alison Sondhaus Carroll
Date:	October 9, 2001
Re:	certification

International patent application No. PCT/FR95/00520
International filing date: April 20, 1995
Applicant: Rhone-Poulenc Rorer, S.A.

Title: "Biologically Active Polypeptides Inserted into an Albumin"

Ref no.: 06832.0004-02000
Translated from the French by Commonwealth Language Services, Ltd.

PCT

World Intellectual Property Organization
International Office

International Patent Application published in accordance with the Patent Convention Treaty (PCT)

(51) International patent classification [6]:
C12N 15/62, 15/14. 15/12, 15/50, 15/80,
15/81, 1/19, 1/15, A61K 38/38, 38/17, A1
C07K 14/765 // C12N 9/88, 1/21, 5/10

(11) International Publication No.: WO 95/30759
(43) International Publication Date: 11/16/95

(21) International application no: PCT/FR95/00520

(81) Designated countries: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, IE, IT, LU, MC, NL, PT, SE)[1]

(22) International filing date: 4/ 20/95

Published
With international search report

(30) Information related to priority:
94/05616 5/6/94 FR

(71) Applicant (*for all designated countries except US*):
RHONE-POULENC RORER S.A. (FR/FR); 20,
Avenue Raymond-Aron, F-92160 Antony (FR)

(72) Inventors; and
(75) Inventors/Applicants (*US only*): BECQUART,
Jerome (FR/FR), 8 rue Maublanc, F-75015 Paris (FR);
CONSEILLER, Emmanuel (FR/FR), 10 rue de Plelo,
F-75015 Paris (FR); GUITTON, Jean-Dominique (FR/FR),
74, rue Dunois, F-75013 Paris (FR); HARDY, Florence (FR/FR),
6, square Villaret-de-Joyeuse, F-75017 Paris (FR); YEH,
Patrice (FR/FR), 11 bis rue Lacepede, F-75005 Paris (FR)

(74) Agent: LE COUPANEC, Pascale, Rhone-Poulenc Rorer
S.A., Patent Department, 20, avenue Raymond-Aron, F-92160
Antony (FR)

(54) Title: BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN

(54) [title in French]

(57) Abstract

Biologically active recombinant polypeptides essentially consisting of at least one active portion derived from a natural or artificial biologically active peptide and inserted into an albumin or albumin variant, the preparation thereof, and pharmaceutical compositions containing same, are disclosed.

(57) [abstract in French]

[1] AT = Austria, BE = Belgium, CH = Switzerland, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = United Kingdom, IE = Ireland, IT = Italy, LU = Luxemburg, MC = Monaco, NL = Netherlands, PT = Portugal, SE = Sweden

BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN

The present invention relates to new biologically active polypeptides, their preparation and pharmaceutical compositions containing them.

Although possessing one or several potential therapeutic activities, numerous polypeptides unfortunately cannot be exploited pharmaceutically. This may be for different reasons, in particular their weak stability *in vivo*, their complex or fragile structure, the difficulty of producing them on an acceptable industrial scale, etc. Likewise, certain polypeptides do not give the expected results *in vivo*, due to problems of administration, conditioning, pharmacokinetics, etc.

The aim of the present invention is precisely to overcome these drawbacks.

It is particularly aimed at the development of artificial, biologically active proteins permitting optimal exploitation of the biological properties of these polypeptides for therapeutic use.

The applicant accordingly provides evidence that using genetic methods, it is possible to insert any active structure derived from a biologically active protein into another protein structure consisting of albumin, without altering the said biological properties. In an unexpected way, human serum albumin allows the active structure to be efficiently introduced at its sites of interaction, and increases the plasmatic stability of the recombinant polypeptide embodied in the invention.

More precisely, the present invention relates to a recombinant polypeptide containing at least one active portion derived from a natural or artificial biologically active polypeptide inserted genetically into an albumin or one of its variants or derivatives.

According to the present invention, a variant of albumin is understood to designate any protein with long plasma half-life obtained by modification of a gene coding for a given isomorph of the human serum albumin by genetic

engineering techniques. It is also understood to designate any macromolecule with long plasma half-life obtained by *in vitro* modification of the protein coded for by such genes. (Modification is understood to mean any mutation, substitution, deletion, addition or modification of a genetic and/or chemical nature.) Albumin being very polymorphic, numerous natural variants have been identified and listed [Weitkamp L.R. et al., Ann. Hum. Genet. 37 (1973) 219].

Derivatives of albumin most particularly refers to molecules consisting wholly or partly of albumin, linked with at least one polypeptide sequence resulting from a natural or artificial gene, whether or not it has biological activity itself.

In the description that follows, the different types of albumins clarified above are commonly designated by the term "albumin."

In the sense of the present invention, active portion is understood to mean a portion having an activity which can be either direct (treatment of diseases, diagnosis, biological research, markers) or indirect (for example, usable in preventing diseases, in the design of vaccines, in medical imaging techniques, etc).

The active portions of biologically active polypeptides inserted according to the invention are preferably of therapeutic interest.

Polypeptides having therapeutic activity may or may not be of human origin.

As representative of polypeptides of non-human origin, one may cite peptides or their derivatives having properties potentially useful in the pathologies of blood or interstitial compartments, such as hirudine, trigramine, antistatine, the anticoagulant peptides of ticks (TAP), arietine, applagine, etc.

According to a preferred embodiment of the invention, the polypeptide having a therapeutic activity is a polypeptide of human origin or a molecular variant. For example, it may consist, wholly or partly, of an enzyme, an enzyme inhibitor, an antigen, an antibody, a hormone, a receptor, a factor

involved in the control of coagulation, an interferon, a cytokine [the interleukins, but also the natural variant antagonists of their attachments to receptors, the SIS (small induced secreted)-type cytokines and, for example, the inflammatory proteins of macrophages (MIPs), etc], a growth and/or differentiation factor [and for example the transforming growth factors (TGFs), the differentiation factors of blood cells (erythropoietin, M-CSF, G-CSF, GM-CSF, etc.), insulin and the growth factors that resemble it (IGFs), or also the cellular permeability factors (VPF/VEGF), etc.], a factor implicated in the genesis or resorption of bone tissue (OIF and osteospontine, for example), a factor implicated in cellular motility or migration [for example, the autocrine mobility factor (AMF), the migration stimulating factor (MSF), or also the dispersion factor (scatter factor/hepatocyte growth factor)], a bactericidal or antifungal factor, a chemotactic factor [for example, the platelet factor 4 (PF4), or the monocyte chemoattractant peptides (MCP/MCAF) or neutrophils (NCAF), etc.], a cytostatic factor (for example, the proteins that attach to galactosides), an adhesive molecule, either plasmatic (for example, the von Willebrand factor, fibrinogen, etc.), or interstitial (laminine, tenascene, vitronectine, etc.), or extracellular matrices, or also any peptide sequence antagonistic or agonistic to molecular and/or intercellular interactions implicated in the pathologies of circulatory and interstitial compartments, for example the formation of arterial or venous thrombi, cancerous metastases, tumoral angiogenesis, inflammatory shock, autoimmune diseases, bone and osteo-articular pathologies, etc.

Of course, the active portion of the polypeptides of the invention may consist of the entire biologically active peptide or a structure derived from it, or may also correspond to a non-natural peptide sequence, isolated for example from random peptide banks. (For purposes of simplification, these different possibilities will be covered in what follows by the collective designation "active portion of a biologically active peptide.") In the sense of the present invention, derived structure is understood to mean any polypeptide obtained by modification which retains its biological activity. Modification is understood to mean any mutation, substitution, deletion, addition or modification of genetic and/or chemical nature. Such derivatives may be generated for different purposes, especially

such as increasing the affinity of the molecule for its attachment sites, improving its level of production, increasing its resistance to proteases, enhancing its therapeutic efficacy or reducing its secondary effects, or giving it new biological properties. As an example, the chimeric polypeptides of the invention have pharmacokinetic properties and biological activity usable in the prevention or treatment of diseases.

The polypeptides according to the invention that are particularly advantageous are those in which the active portion has:

(a) the entire peptide structure, or
(b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or several residues) and possessing therapeutic activity.

Among the structures of type (b) may be cited in particular the molecules in which certain sites of N- or O-glycosylation have been modified or deleted, molecules in which one or several residues have been substituted, or molecules in which all the cysteine residues have been substituted. One may also cite molecules obtained from (a) by deleting the regions having little or no involvement in the interaction with the linkage sites under consideration, or expressing an undesirable activity, and molecules including supplementary residues with respect to (a), for example, a terminal methionine-N, and/or a secretion signal, and/or a junction peptide.

The subject matter of the present invention is particularly advantageous in regard to active peptides that are too small to form a structural domain, and/or do not possess good stability *in vivo* and/or good bioavailability. The proposed method of insertion according to the invention allows them to be associated with one or more preexisting domains of albumin and thus to benefit from the bioavailability and stability *in vivo* of the latter.

In general, the sizes of the active portions inserted into the albumin vary between three and twenty five amino acid residues. However, sequences of 1 to 100 residues can also be used.

The insertion of an active portion of a peptide into the peptide sequence of albumin is accomplished according to the invention in a way satisfying the following two conditions:

Sufficient accessibility must be preserved to keep intact the biological activity of the active portion inserted into the albumin. Moreover, the structure of the albumin of course cannot be subjected to a destabilization significant enough to be detrimental to the recombinant polypeptide, or chimera.

The sites of insertion are preferably selected within the albumin while adhering to the above requirements.

According to the crystalline structure published by He and Carter (Nature 1992, 358, 209-215), albumin is formed by the repetition of 3 domains, each comprising two subdomains, and it consists of over 67% alpha helices. Each domain is superimposable on the others and is composed of 10 helices designated h1 to h10. The subdomain A consists of helices h1 to h6 and the subdomain B, helices h7 to h10. Each subdomain is formed in a common pattern: h1, h2, h3, h4 for domain A, and h7, h8, h9 and h10 for domain B. The small supplementary helices h5 and h6 are linked by a disulfide bond to subdomain A. Figure 1 schematizes the structure of human serum albumin.

The sites of insertion are preferably located in the regions of the albumin presumably forming the exposed regions on the surface of the molecule; these regions preferentially being loops.

As insertion sites that are particularly suitable for the invention, three regions of the first domain may be mentioned:

- region 5, which extends from residue 57 to 62 and corresponds to a loop joining the helices h3 and h4;
- region 8, comprising the residues 103 to 120 and corresponding to the inter-subdomain zone;
- region 13, between residues 178 and 200, which corresponds to a helix.

According to the invention, helices h2 and h3 of domain III delimit another conceivable insertion region, from residues 419 to 430.

An active portion of the biologically active peptide may be inserted in the albumin peptide sequence according to three different embodiments of the invention:

- A strict insertion consisting of a simple addition of the peptide sequence of interest into the original sequence of the albumin, which is totally conserved.
- The insertion may correspond to a substitution of a portion of the albumin peptide sequence by the peptide sequence corresponding to the active portion of the peptide of interest.
- Finally, an insertion combining an addition of a part of the active peptide sequence and a substitution of a portion of the albumin peptide sequence by the rest of the active portion of the active peptide.

Figure 2 schematically depicts these different modes of insertion.

Of course, the active portion of a biologically active peptide can be repeated several times in the chimera at the same place and/or in different regions of the albumin. Likewise, it is equally possible as another embodiment of the invention to insert different active portions, whether stemming from the same peptide or from different peptides.

Finally, in another embodiment of the invention, an active portion of a peptide may be inserted either within the albumin or by surrounding the sequences of the junction.

In regard to the junction sequences, peptide sequences rich in glycine residues and/or serine residues and/or threonine residues, and/or any amino acid residue described as frequently encountered in the zones of flexibility in proteins, may be especially involved.

The recombinant polypeptides according to the invention prove to be particularly advantageous.

They permit a given biological activity to be maintained in an organism for an extended period. It thus proves to be possible to reduce the administered doses, and in certain cases to potentiate the therapeutic effect, for example by reducing the consecutive secondary effects of a larger dose. Advantageously, they allow the production and utilization of structures derived from biologically active polypeptides that are very small and therefore very specific to a desired effect. Moreover, these recombinant polypeptides are distributed in the organism especially advantageously,

modifying their pharmacokinetic properties and favoring the development of their biological activity and their utilization. They also have the advantage of

being weakly or nonimmunogenic for the organism in which they are used. Finally, the polypeptides of the invention may be expressed (or preferably secreted) by recombinant organisms at levels permitting their industrial use.

Another aim of the invention relates to a process for preparing the chimeric molecules described above. More precisely, this process consists of causing a nucleotide sequence coding for an active portion of the desired polypeptide to be expressed by a eukaryote or prokaryote cellular host, then harvesting the product polypeptide.

Among the eukaryote hosts usable within the scope of the present invention one may cite animal cells, yeasts or fungi. In particular, for yeasts one may cite those of the type Saccharomyces, Kluyveromyces, Pichia, Schwanniomyces or Hansenula. For animal cells, COS, CHO, C127, etc. may be cited. Among the fungi able to be used according to the present invention one may cite in particular Aspergillus ssp, or Trichoderma ssp. As prokaryote hosts, bacteria such as Escherichia coli or those belonging to the type Corynebacterium, Bacillus or Streptomyces would be preferable.

The nucleotide sequences usable within the scope of the present invention can be prepared in different ways. Generally they are obtained by assembling in the reading frame the sequences coding for each of the functional parts of the polypeptide. These can be isolated by techniques of a person skilled in the art, for example, directly from cellular ARN messengers (ARNm), or by recloning from a bank of complementary ADN (ADNc), or again, totally synthetic nucleotide sequences may be involved. Moreover, it is understood that nucleotide sequences may also be subsequently modified, for example, by genetic engineering techniques, to obtain derivatives or variants of the said sequences.

The insertion of this nucleotide sequence coding for the active portion of the polypeptide may or may not be directly achieved, depending on the region chosen for the insertion site in the gene coding for albumin.

The region selected may not actually include the restriction site adequate for the said insertion. According to this hypothesis, it may prove to be of interest to introduce one or several single restriction sites prior to the insertion.

The creation of restriction sites at the level of the region chosen for the insertion site is preferably done by directed mutagenesis according to traditional techniques. However, one can also insert the nucleotide sequence corresponding to "the active portion of the biologically active peptide" directly by directed mutagenesis so that the insertion does not cause particular restriction sites to appear.

In the particular case of region 5 in the gene coding for albumin, the presence of restriction site PvuII permits the direct insertion of the nucleotide sequence. It thus shows itself to be particularly useful as a cloning site for an active portion of a polypeptide which one hopes to insert into the translational [reading] frame within the albumin sequence at the 57^{th} residue level.

In this mode of insertion, which makes use of a single restriction site already existing in the original albumin sequence, the ligation of the sequence coding for the active peptide with the restriction fragment corresponding to the totality of the gene coding for albumin, generates a nucleotide sequence comprising a hybrid gene coding for a strict insertion HAS-type chimeric protein.

In the particular case of zones 419 to 430, the presence of 2 insertion sites (HincII and ArvII) allows the peptides of biological interest to be inserted and/or substituted in the albumin sequence.

With particular reference to regions 8 and 13, the insertion of the said sequence is favored if one or more restriction sites that can be manipulated are created for it beforehand. To be sure, the restriction sites to be created on the sequence level are chosen by taking into account the nature of the preexisting restriction sites. They must lead to a selective insertion.

In the second embodiment, the use of two single restriction sites allows the construction of genes coding for chimeras having the active peptide inserted and/or substituted. The peptide insertion is made by replacement of a narrow fragment of two single restriction sites in the complementary ADN of the albumin, inserted by directed mutagenesis. These two single restriction sites can be Mst I and Kpn I in region 8 and Sst I and Xho I in region 13. The creation of these restriction sites may or may not modify the polypeptide

sequence of human albumin. The subsequent cloning of active peptide in coding phase in this albumin gene may be exclusively the peptide

coding sequence, or may correspond to a blend of the coding sequence of the peptide and the coding sequence of the deleted albumin fragment.

The present invention also aims to protect the variants of the nucleotide sequences coding for the corresponding albumin; that is, integrating at least one single non natural restriction site, not present in the original sequence.

Creation of these restriction sites can additionally serve to advantage in inserting one or several active portions of biologically active polypeptides into the mature protein.

More preferably, in the process that is the subject matter of the invention, the nucleotide sequence is part of an expression cassette comprising a transcription initiation region (promoter region), permitting expression of the nucleotide sequence in the host cells placed under its control and coding for the polypeptides in the invention. This region can arise from promoter regions of genes strongly expressed in the cellular host utilized, the expression being constitutive or regulated. With respect to yeasts, it may be the promoter of the gene for phosphoglycerate kinase (PKG); for glyceraldehyde-3-phosphate dehydrogenase (GPD); lactase (LAC4); the enolases (ENO); the alcohol dehydrogenases (AOH), etc. As to bacteria, it may be the promoter of the right or left genes of the lambda bacteriophage (P_L, P_R), or also the promoters of genes of the operons tryptophane (P_{trp}) or lactose (P_{lac}). Moreover, this control region can be modified, for example by mutagenesis *in vitro*, by introduction of additional control elements or synthetic sequences, or by deletions or substitutions of the original control elements. The expression cassette can equally comprise a region of functional transcription termination in the host considered, positioned immediately ahead of the nucleotide sequence coding for a polypeptide embodied in the invention.

In a preferred embodiment, the polypeptides of the invention result from the expression of a nucleotide sequence in a eukaryote or prokaryote host and the secretion of the expression product of the said sequence into the culture medium. In fact, it is particularly advantageous to be able to obtain the molecules by recombinant pathway directly in the culture medium. In this case, the nucleotide sequence coding for a polypeptide embodied in

the invention is preceded by a "leader" sequence (or signal sequence) directing the polypeptide arising in the secretion pathways of the host utilized. This "leader" sequence may be the natural signal sequence of the biologically active peptide in the case where this is a naturally secreted protein, or may be that of the stabilizing structure, but it may equally well consist of any other functional "leader" sequence, or of an artificial "leader" sequence. The choice of one or another of these sequences is guided in particular by the host utilized. Examples of functional signal sequences include those of the sexual pheromone genes or the "killer" toxins of yeasts.

Besides the expression cassette, one or more markers permitting the selection of the recombinant host can be added, such as for example the gene URA3 of the yeast S.Cerevisiae, or the genes conferring resistance to antibiotics such as geneticine (G418), or any other toxic compound such as certain metallic ions.

The ensemble comprising the expression cassette and the selection marker may be introduced directly into the host cells in question, that is, inserted beforehand in a functionally self-replicating vector. In the first case, sequences homologous to the regions present in the genome of the host cells are preferably added to this ensemble; the said sequences being then positioned on each side of the expression cassette and of the selection gene as a way to increase the integration frequency of the ensemble in the host genome by targeting the integration of the sequences by homologous recombination. In the case where the expression cassette is inserted into a replicative system, a replication system preferred for yeasts of the type Kluyveromyces is derived from the plasmid pKD1, initially isolated from K. drosophilianum. A replication system preferred for the yeasts of type Saccharomyces is derived from plasmid 2μ of S.Cerevisiae. Furthermore, this expression plasmid may contain all or part of the said replication systems, or may combine elements derived from plasmid pkD1 as well as from plasmid 2μ.

In addition, the expression plasmids may be shuttle vectors between a bacterial host such as Escherichia coli and the chosen cellular host. In this case, a replication origin and a selection marker functioning in the bacterial host are required. It is also possible to position the restriction sites of

around the bactérial sequences and only on the expression vector; this makes it possible to delete [these sequences] by cutting and religation *in vitro* of the truncated vector before transformation of the host cells, which may result in an increased number of copies and an enhanced stability of the expression plasmids in the said hosts. For example, these restriction sites may correspond to sequences such as 5'-GGCCNNNNNGGCC-3' (Sfil) (SEQ ID N°1) or 5'-GCGGCCGC-3' (NotI) (SEQ ID N°2), inasmuch as these sites are extremely rare and generally absent from an expression vector.

After construction of such vectors or expression cassettes, they are introduced into the host cells retained according to the traditional techniques described in the literature. In this regard, any method permitting the introduction of a foreign ADN into a cell may be used. This includes in particular transformation, electroporation, conjugation or any other technique known to a person skilled in the art. As an example for the hosts of yeast type, the different strains of Kluyveromyces utilized have been transformed by treating the intact cells in the presence of lithium acetate and polyethylene glycol, according to the technique described by Ito et al. [J. Bacteriol. 153 (1983) 183]. The transformation technique described by Durrens et al. [Curr. Genet. 18 (1990) 7] utilizing ethylene glycol and dimethylsulfoxide has also been used. It is also possible to transform yeasts by electroporation, according to the method described by Karube et al. [FEBS Letters 182 (1985) 90]. An alternative protocol is also described in detail in the examples that follow.

After the selection of transformed cells, the cells expressing the said polypeptides are inoculated and the said polypeptides may be recovered, either during cellular growth for the "ongoing" processes, or at the end of growth for batch cultures. The polypeptides that are the subject matter of the present invention are purified using the supernatant of the culture, with the aim of determining their molecular, pharmacokinetic and biological nature.

A preferred system of expression of the polypeptides according to the invention consists of using the yeasts of the Kluyveromyces type as cellular hosts, transformed by certain vectors derived from the extrachromosomic replicon pKD1 initially isolated from K.marxianus var. drosophilianum.

These yeasts, in particular K.lactis and K.Fragilis, are generally capable of replicating the said vectors in a stable manner and furthermore have the

advantage of being included in the G.R.A.S. (Generally Recognized As Safe) list of organisms. Preferred yeasts are preferably the industrial strains of the Kluyveromyces type capable of replicating in stable fashion the said plasmids derived from plasmid pKD1 and in which a selection marker and expression cassette have been inserted permitting the secretion of a high level of the polypeptides according to the invention.

The present invention also relates to the nucleotide sequences coding for the chimeric polypeptides described earlier, as well as the recombinant cells, eukaryotes or prokaryotes, comprising such sequences.

The present invention also relates to the application of the polypeptides according to the present invention as medicines. More particularly, the aim of the invention is any pharmaceutical composition comprising one or several polypeptides or nucleotide sequences such as are described earlier. These pharmaceutical compositions can be presented in various formulations. They may particularly involve nanoparticles having surfaces on which the polypeptides according to the invention are present. This type of formulation is utilized more particularly to achieve directed targeting of the active [molecule]. Of course, the nucleotide sequences can be used in gene therapy.

The present invention will be described more completely with the aid of the examples that follow, which should be considered as illustrative and not limiting.

LIST OF FIGURES

The representations of plasmids indicated in the following figures are not drawn to scale and only the restriction sites that are important for understanding the clonings carried out are indicated.

Figure 1: Schematic representation of the domain I of albumin with location of the insertion sites according to the invention. * signals the location of the h2-h3 insertion site in domain III. The numbers 5, 8 and 13 identify the corresponding insertion zones.

Figure 2: Schematic representation of different modes of insertion of an active peptide in the structure of albumin.

Figure 3: Plasmid pYG105.

Figure 4: Modification of region 5 of albumin following the insertion of the sequence coding for IEGR, as described in example 8.1 The modifications are represented in bold type. The placement of a modified or introduced restriction site is indicated by a horizontal line, and the position of an enzyme cutting with a vertical line.

Figure 5: Cloning strategy for insertion of IEGR in region 5 (example 8.1).

Figure 6: Cloning strategy for insertion of an active peptide in region 13 of albumin.

EXAMPLES

GENERAL CLONING TECHNIQUES

The traditional methods used in molecular biology, such as the preparative extractions of plasma ADN, centrifugation of plasma ADN in a cesium chloride gradient, electrophoresis on agar gels or acrylamide, purification of ADN fragments by electroelution, extraction of proteins in phenol or phenol-chloroform, precipitation of ADN in saline medium by ethanol or isopropanol, transformation in Escherichia coli, etc., are well known to a person skilled in the art and are abundantly described in the literature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982; Ausubel F.M. et al. (eds), "Current Protocols in Molecular Biology," John Wiley & Sons, New York, 1987.]

Restriction enzymes were supplied by New England Biolabs (Biolabs), Bethesda Research Laboratories (BRL) or Amersham, and are used according to the recommendations of the suppliers.

Plasmids of type pBR322, pUC and the M13 series phages are commercial in origin (Bethesda Research Laboratories).

For the ligations, the ADN fragments are separated according to size by electrophoresis on agar gel or acrylamide, extracted with phenol or a

phenol/chloroform mixture, precipitated with ethanol and incubated in the presence of the ADN ligase of phage T4 (Biolabs) according to the supplier's recommendations.

The filling of prominent 5' extremities is achieved by the Klenow fragment of E. coli ADN polymerase (Biolabs), used according to the supplier's specifications. The prominent 3' extremities are destroyed in the presence of the ADN polymerase of T4 phage (Biolabs) according to the manufacturer's recommendations. The prominent 5' extremities are destroyed by a treatment brought about with S1 nuclease.

Directed mutagenesis *in vivo* with synthetic oligodeoxynucleotides is achieved according to the method developed by Taylor et al. [Nucleic Acids Res. 13 (1985) 8749-8784], using the kit distributed by Amersham. In the context of the present invention, this technique is utilized in particular to create single restriction sites for use in a subsequent insertion.

Enzymatic amplification of ADN fragments by the technique called PCR [Polymerase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. and Faloona F.A., Math. Enzym. 155 (1987) 335-350] is carried out with the use of a "DNA thermal cycler" (Perkin Elmer Cetus) according to the manufacturer's specifications.

The nucleotide sequences are verified using the method developed by Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5487], using the kit distributed by Amersham.

K.lactis is transformed using the ADN from the expression plasmids of the proteins of the present invention by any technique known to a person skilled in the art. An example is given in the text.

Except where otherwise indicated, the bacterial strains used are E.coli MC1060 (laclPOZYA, X74, galU, galK, $strA^{1}$), or E.coli TG1 (lac, proA,B, supE, thi, hsdD5/ F'traD36, $proA^{+}B^{+}$, $lacI^{Q}$, lacZ, M15).

The yeast strains used are those of budding yeast, in particular yeasts of the Kluyveromyces type. The K. lactis MW98-8C strain (a. uraA, aro, lys, K^{+}, $pKD1^{e}$) and K. lactis CBS 293.91 were used in particular. A sample of the

MW98-8C strain was deposited on September 16, 1988 at Centraalbureau voor Schimmelkulturen (CBS), Baarn (Netherlands) where it was registered under the number CBS 579.88.

A bacterial strain (E. coli) transformed with the plasmid pET-8c52K was deposited on April 17, 1990 with the American Type Culture Collection under the number ATCC 68306.

Yeast strains transformed by the expression plasmids coding for the proteins of the present invention are cultivated in Erlenmeyer flasks or 2 l pilot fermenters (SETRIC, France), in rich medium (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; or YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) under constant stirring at 28°C.

EXAMPLE 1:
PROTOCOL FOR STRICT PEPTIDE INSERTION UTILIZING A SINGLE RESTRICTION SITE PRESENT IN THE HSA GENE.

Due to its uniqueness in the HSA gene and the associated vector, the site PvuII, located naturally in the coding sequence, is particularly useful as a cloning site for a biologically active peptide that one wishes to insert in translation phase into the HSA at the level of the 58th residue. In a particular embodiment, it is useful to employ peptides from p residues for which the coding sequence is $[3xN]_p$. In this case, the oligonucleotides synthesized are of type 5'-NN$[3xN]_p$N-3' and its complementary strand. The ligation of this fragment with the restriction fragment HindIII-HindIII Δ PvuII, corresponding to the totality of the gene coding for the HSA, generates a HindIII-HindIII restriction fragment comprising a hybrid gene coding for a chimeric protein of strict insertion type HSA. In another embodiment, the peptide may be repeated several times in the chimera.

EXAMPLE 2:
TOTAL OR PARTIAL PEPTIDE INSERTION PROTOCOL UTILIZING TWO SINGLE AND NATURAL RESTRICTION SITES.

In a particular embodiment, the use of two single restriction sites allows for the creation of genes coding for chimeras having the active peptide in strict insertion or partial insertion.

For example, the existence of single HincII and AvrII sites in the coding sequence of HSA and in the vector allows an HindIII-HindIII Δ HincII-AvrII fragment to be generated. The elimination of the nucleotide fragment HincII-AvrII corresponds to the deletion of the peptide fragment T(420) – N(429). The use of an appropriate complementary oligonucleotide permits a peptide in cloning phase to be cloned in the HSA gene. This restriction fragment may be exclusively the complementary coding sequence of the peptide, or may correspond to a blend of the coding sequence of the active peptide and the coding sequence of the T(420) – N(429) fragment of the HSA. The active peptide may be present several times in the chimera.

EXAMPLE 3
EXPRESSION PLASMIDS

The chimeric proteins of the preceding examples may be expressed in the yeast, starting from functional promoters, regulated or constitutive, such as, for example, those present in the plasmids pYG105 (promoter LAC4 of Kluyveromyces lactis), pYG106 (promoter PGK of Saccharomyces cerevisiae), pYG536 (promoter PHO5 of S. cerevisiae), or the hybrid promoters such as those described in the patent application EP 361 991. The plasmids pYG105 and pYG106 are particularly useful here because they permit the expression of genes coded by the restriction fragments HindIII that are described in the preceding examples and are cloned in the HindIII site and in productive orientation (defined as the orientation that places the "prepro" region of the albumin in proximal fashion relative to the transcription promoter), from the functional promoters in K. lactis, whether regulatable (pYG105), or constituent (pYG106). The plasmid pYG105 corresponds to the plasmid pKan707 described in patent application EP 361 991, in which the restriction site HindIII, single and located in the gene for resistance to geneticine (G418), has been destroyed by directed mutagenesis while preserving a protein unchanged (oligodeoxynucleotide 5'-GAAATGCATAAGCTCTTGCCATTCTCACCG-3' (SEQ ID No. 3). The fragment SalI-SacI coding for the gene URA3 of the transferred plasmid was replaced by a restriction fragment SalI-SacI, comprising an expression cassette consisting of the promoter LAC4 of K. lactis (in the form of a fragment SalI-HindIII) and the terminator of the PGK gene of S. cerevisiae (in the form of a fragment HindIII-SacI). The plasmid pYG105 is very stable mitotically in the Kluyveromyces yeasts. It

is represented in Figure 3. The plasmids pYG105 and pYG106 do not differ from each other than in the nature of the transcription promoter encoded by the fragment SalI-HindIII.

EXAMPLE 4:
TRANSFORMATION OF YEASTS

The transformation of yeasts belonging to the Kluyveromyces type, and particularly the strains MW98-8C and CBS 293.91 of K. lactis, is carried out, for example, by the technique of treating the intact cells with lithium acetatye [Ito H. et al., J. Bacteriol. 153 (1983) 163-168], adapted as follows. The growth of the cells takes place at 28°C in 50 ml of YPD medium through stirring and up to an optical density between 0.6 and 0.8 at 600 nm (OD_{600}); the cells are harvested by centrifugation at low speed, washed in a sterile solution of TE (10 mM Tris HCl pH 7.4; 1 mM EDTA), resuspended in 3–4 ml of lithium acetate (0.1 M in the TE) to obtain a cellular density of about 2 x 10^8 cells/ml, then incubated at 30°C for 1 hour with moderate stirring. Aliquots of 0.1 ml of the resulting suspension of competent cells are incubated at 30°C for 1 hour in the presence of ADN and at a final concentration of 35% polyethylene glycol (PEG_{4000}, Sigma). After a thermal shock of 5 minutes at 42°C, the cells are washed twice, resuspended in 0.2 ml of sterile water and incubated for 16 hours at 28°C in 2 ml of YPD medium to permit the phenotypic expression of the gene conferring resistance to G418, expressed under control of the promoter P_{kl} (cf. EP 361 991); 200 µl of the cellular suspension is then spread over YPD selective containers (G418, 200µg/ml). The containers are put into incubation at 28°C and the transformed cells appear after 2 to 3 hours of cellular growth.

EXAMPLE 5:
SECRETION OF CHIMERAS

After selection in rich medium supplemented with G418, the recombinant clones are tested for their capacity to secrete the mature form of chimeric proteins. Clones corresponding to the CBS 293.91 or MW98-8C strain, transformed by the expression plasmids of the chimeras between the HSA and the biologically active portion are incubated in YPD or YPL medium at 28°C. Cells in the supernatant are retrieved by centrifugation when the cells reach stationary growth phase, concentrated tenfold if necessary by precipitation for 30 minutes at 20°C in a final concentration of 60% ethanol, and tested after electrophoresis on SDS-PAGE 0 gel, either directly by coloration of the gel by coomassie blue, or after

immunoblot, using primary antibodies directed against the biologically active portion or a polyclonal rabbit serum directed against the HSA. During the immunologic detection experiments, the nitrocellulose filter is first incubated in the presence of the specific primary antibody, washed several times, incubated in the presence of goat antibodies directed against the primary antibody, then incubated in the presence of an avidine-peroxidase complex utilizing the "ABC kit" distributed by Vectastain (Biosys S.A., Compiegne, France). The immunologic reaction is then revealed by addition of diamino-3,3' benzidine tetrachlorhydrate (Prolabo) in the presence of oxygenated water, according to the manufacturer's recommendations.

EXAMPLE 6;
PURIFICATION OF CHIMERAS

The chimeras present in the supernatants of the cultures corresponding to the transformed strain CBS 293.91 are characterized at a first stage with the aid of antibodies specific to the HSA part. It may be desirable to purify some of these chimeras. The culture is then centrifuged (10,000 g, 30 min), the supernatant is passed through a 0.22 mm filter (Millipore), then concentrated by ultrafiltration (Amicon) using a membrane for which the discrimination threshold is at 30 kDa. The concentrate obtained is then dialyzed using a solution of Tris HCl (50 mM, pH 8) and purified by column. For example, the concentrate corresponding to the supernatant of the culture from the transformed strain CBS 293.91 is purified by affinity chromatography on Trisacryl Blue (IBF), and the samples are next dialyzed using water. Purification by molecular sieve may then be carried out. In this case, a Superose 12 column (Pharmacia) is equilibrated beforehand with a buffer of 20 mM NaH_2PO_4 and 100 mM NaCl at pH 7.0, and the samples resulting from the affinity chromatography are loaded on the column, harvested and characterized. Purification by ion exchange chromatography may also be used. In certain cases, the concentrate obtained after ultrafiltration is dialyzed using a solution of Tris HCl (50 mM, pH 8), then deposited in 20 ml fractions on a cation exchange column (5 ml) (S Fast Flow, Pharmacia) equilibrated with the same buffer. The column is then washed several times with the Tris HCl solution (50 mM, pH 8), and the chimera protein is then eluted in the column by an NaCl gradient (0 to 1 M).

The fractions containing the chimera protein are then combined and dialyzed using a solution of Tris HCl 50mM (pH 8) and redeposited on the S Fast Flow column. After elution of the column, the fractions containing the protein are combined, dialyzed using water and lyophilized before characterization.

EXAMPLE 7:
INSERTION OF PEPTIDE 11 IN THE ALBUMIN SEQUENCE

Peptide 11 was described as an epitote of tryptophan synthase (Larvor et al., Mol. Immunol. (1991) 28, 523-531).

1. Strict insertion in a region naturally having single restriction sites

Peptide 11 has the following peptide sequence in the terminal N a C sense: HGRVGIYFGMK (SEQ ID No.20). The strategy of insertion in the HincII-AvrII zone has consisted of creating a ligation between a synthesized nucleotide fragment coding for peptide 11 such that the reading frame for the sequence coding for HSA is respected, and the nature of the sequence coding for HSA is as it originally was. In a first stage we have thus synthesized the following two oligonucleotides: 5'-CCATGGTAGAGTAGGTATCTATTTCGGTATGAAAACTCCAACT CTTGTAGAGGTCTCGAGAAAT-3' (SEQ ID No.4) and 5'-CTAGATTTCTCGAGACCTCTACAAGATGTGGAGTTTTCATACCGA AATAGATACCTACTCTACCATGG-3' (SEQ ID No.5).
These two oligonucleotides have been hybridized and then ligated from the fragment HindIII-HindIII Δ HincII-AvrII, thus generating the totality of the HSA gene comprising the total insertion of peptide 11 between S(419) and T(420), immediately preceding the "prepro" exportation region of the HSA. This fragment is cloned in productive orientation and in the HindIII site of the pYG105 plasmid.

2. Insertion by substitution and addition in a region naturally having single restriction sites

In another embodiment, the oligonucleotides synthesized were the following: 5'-CCATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3' (SEQ ID No.6), and
5'- CTAGTTTCATACCGAAATAGATACCTACTTCTTACCATGG-3' (SEQ ID No.7). These two oligonucleotides are hybridized and then ligated to the fragment HindIII-HindIII Δ HincII-AvrII, thus generating the gene of a chimera severed from the residue T(420) to N(429) and substituted by the sequence of peptide 11. This fragment is cloned in productive orientation and in the HindIII site of the pYG105 plasmid.

3. Strict insertion in a region not having single restriction sites

In another embodiment, two single restriction sites, Sst I and Xho 1, have been created by directed mutagenesis. The same type of directional cloning was achieved. In the case of total insertion of peptide 11 between residue A(191) and S(192), the following two oligonucleotides were synthesized:
5'-ACGGGATGAAGGGAAGGCCCATGGTAGAGTAGGTATCTATTT CGGTATGAAA-3' (SEQ ID No.8), and
5'-TCGATTTCATTACCGAAATAGATACCTACTCTACCATGGGCCTT CCCTTCATCCCGTAGCT-3' (SEQ ID No.9).
One then proceeds according to the protocols already described in the preceding points 1 or 2 to obtain the corresponding expression plasmid.

4. Insertion by substitution in a region not having single restriction sites

The insertion requires the preliminary creation of two restriction sites, SstI and XhoI. For the partial insertion itself, the following two oligonucleotides were synthesized:
5'-ACATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3'
(SEQ ID No.10), and
5'-TCGATTTCATACCGAAATAGATACCTACTCTACCATGTAGCT-3'
(SEQ ID No.11).
In this last case, the residues R(186) to A(191) are deleted and substituted by the peptide 11. The two expression plasmids of these chimeras are 1671 and 1667, respectively.

Each of the plasmids obtained in points 1, 2, 3 and 4 is used to transform a yeast strain, according to the protocol described in example 4. The corresponding proteins are secreted and purified as in examples 5 and 6.

EXAMPLE 8:
INSERTION OF PEPTIDE SEQUENCE IEGR, THE SUBSTRATE OF FACTOR Xa

The insertion of the peptide containing the sequence IEGR (SEQ ID No.21) comes next, target of the protease factor Xa, which turns prothrombin into thrombin in the cascade of reactions involved in the coagulation of blood.

1. Strict insertion in region 5 having a single PvuII site.

To achieve this insertion at the PvuII site of the albumin gene, a first step is to create a replicative vector lacking a PvuII site, in which the gene coding for the prepro albumin is then inserted. The following two complementary oligonucleotides, coding for the sequence IEGR, have also been synthesized:
5'- GATCCATAGAAGGTCGACTAG-3' (SEQ ID No.12), and
3'- CTAGGTATCTTCCAGCTGATC-5' (SEQ ID No.13).
These two oligonucleotides are then hybridized and inserted at the level of site PvuII in the albumin gene. The modifications in the nucleotide and peptide sequences of the albumin that their insertion causes to appear are presented in Figure 4. The cloning strategy is schematized in Figure 5.
This construction illustrates the case where the junction sequences are introduced on either side of the peptide at issue.

2. Insertion in region 13.

Figure 6 depicts a cloning strategy for a peptide in region 13.
In what follows, the restriction sites Sst I and Xho 1 were created by directed mutagenesis. In the case of region 13, a total substitution and an addition substitution of the IEGR sequence were carried out respectively, either simple or surrounded by the junction sequences.

In regard to the total substitution, the oligonucleotides utilized are the following:
5'-CAGAATCGAAGGTAGAGCC-3' (SEQ ID No.14), and
5'-TCGAGGCTCTACCTTCGATCGAGGGTAGCT-3' (SEQ ID NO.15).
On the protein level, the sequence (187)DEGK (SEQ ID No.22) is substituted by IEGR.

In the second case, the oligonucleotides utilized are the following:
5'-ACCCTCGATCGAAGGTAGATCTCCA-3' (SEQ ID No.16)
5'-TCGATGGAGATCTACCTTCGATCGAGGGTAGCT-3'
(SEQ ID No.17)
On the protein level, the albumin is severed from residue R(186) to residue A(191) and replaced by the sequence PSIEGRSP (SEQ ID No.23), thus leading to an addition of two residues.
The corresponding proteins are secreted and purified according to the protocols described in the preceding examples. Table 1 below lists the characteristics of these chimeras.

Table 1

Albumin Incorporating	Expression Rate µg/l	Purification Yield	Final Purity
IEGR	150	15	98
PSIEGRPS	110	36	96

3. Biological activity of the chimera obtained according to example 8.1

The chimera is incubated for 3 hours at 37°C in the presence of bovine factor Xa in an enzyme/ substrate ratio of 1/10 and a buffer of Tris 50mM, NaCl 100mM, and $CaCl_2$ 1mM, at pH 8.0. At the conclusion of this treatment, an SDS PAGE analysis is performed that provides evidence for cleavage of the chimera, characterized by the generation of an albumin fragment with a mass of the order of 80 kD.

LIST OF SEQUENCES

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: RHONE-POULENC RORER S. A.
(B) STREET: 20, avenue Raymond ARON
(C) CITY: ANTONY
(D) COUNTRY: FRANCE
(E) ZIP CODE; 92165

(ii) TITLE OF THE INVENTION: New biologically active polypeptides, their preparation and pharmaceutical compositions containing them.

(iii) NUMBER OF SEQUENCES: 24

(iv) LEGIBLE FORM FOR COMPUTER:
(A) TYPE OF SUPPORT: TAPE
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (OEB)

(2) INFORMATION FOR SEQ ID NO.1:

(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii)ANTI-SENSE: NO

(xi)DESCRIPTION OF SEQUENCE: SEQ ID NO.1

GGCCNNNNNG GCC

(3) INFORMATION FOR SEQ ID NO.2:

(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 8 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc

(iii) HYPOTHETICAL: NO

(iii)ANTI-SENSE: NO

(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.2

GCGGCCGC

(4) INFORMATION FOR SEQ ID NO.3:

(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.3

GAAATGCATA AGCTCTTGCC ATTCTCACCG

(5) INFORMATION FOR SEQ ID NO.4:

(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.4

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAAACTCCA ACTCTTGTAG AGGTCTCGAG AAAT

(6) INFORMATION FOR SEQ ID NO.5:

(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS simple:
(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.5

CTAGATTTCT CGAGACCTCT ACAAGATGTG GAGTTTTCAT ACCGAAATAG ATACCTACTC TACCATGG

(7) INFORMATION FOR SEQ ID NO.6:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.6

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAA

(8) INFORMATION FOR SEQ ID NO.7:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.7

CTAGTTTCAT ACCGAAATAG ATACCTACTT CTTACCATGG

(9) INFORMATION FOR SEQ ID NO.8:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.8

ACGGGATGAA GGGAAGGCCC ATGGTAGAGT AGGTATCTAT TTCGGTATGA AA

(10) INFORMATION FOR THE SEQ ID NO.9:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple

(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO

(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.9

TCGATTTCAT TACCGAAATA GATACCTACT CTACCATGGG CCTTCCCTTC ATCCCGTAGC T

(11) INFORMATION FOR SEQ ID NO.10:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.10

ACATGGTAGA GTAGGTATCT ATTTCGGTAT GAAA

(12) INFORMATION FOR SEQ ID NO.11:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A)LENGTH: 42 base pairs
(B)TYPE: nucleic acid
(C)NUMBER OF STRANDS: simple
(D)CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.11

TCGATTTCAT ACCGAAATAG ATACCTACTC TACCATGTAG CT

(13) INFORMATION FOR SEQ ID NO.12:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A)LENGTH: 21 base pairs
(B)TYPE: nucleic acid
(C)NUMBER OF STRANDS: simple
(D)CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.12

GATCCATAGA AGGTCGACTA G

(14) INFORMATION FOR SEQ ID NO.13:
(i)CHARACTERISTICS OF THE SEQUENCE:

(A)LENGTH: 21 base pairs
(B)TYPE: nucleic acid

(C) NUMBER OF STRANDS: simple
(E) CONFIGURATION: linear
(ii)TYPE OF MOLECULE: ADNc
(iii)HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO.13

CTAGGTATCT TCCAGCTGAT C

(15) INFORMATION FOR SEQ ID NO. 14:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 14

CAGAATCGAA GGTAGAGCC

(16) INFORMATION FOR SEQ ID NO. 15:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 15

TCGAGGCTCT ACCTTCGATC GAGGGTAGCT

17) INFORMATION FOR SEQ ID NO. 16:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 16

ACCCTCGATC GAAGGTAGAT CTCCA

(18) INFORMATION FOR SEQ ID NO. 17:
(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 17

TCGATGGAGA TCTACCTTCG ATCGAGGGTA GCT

19) INFORMATION FOR SEQ ID NO. 18:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 18

GTCCCGGATG GAGCGCGTAC TTAGAGAGAA T

20) INFORMATION FOR SEQ ID NO. 19:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) NUMBER OF STRANDS: simple
(D) CONFIGURATION: linear
(ii) TYPE OF MOLECULE: ADNc
(iii) HYPOTHETICAL: NO
(iii)ANTI-SENSE: NO
(xi)DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 19

TCGACAGGGC CTACCTCGCG CATGAATCTC TCTTAAGCT

21) INFORMATION FOR SEQ ID NO. 20:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) CONFIGURATION: linear
(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 20

His Gly Arg Val Gly Ile Tyr Phe Gly Met Lys

22) INFORMATION FOR SEQ ID NO. 21:
(i) CHARACTERISITCS OF THE SEQUENCE:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 21

Ile Glu Gly Arg

23) INFORMATION FOR SEQ ID NO. 22:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 22

Asp Glu Gly Lys

24) INFORMATION FOR SEQ ID NO. 23:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 23

Pro Ser Ile Glu Gly Arg Ser Pro

25) INFORMATION FOR SEQ ID NO. 24:
(i) CHARACTERISTICS OF THE SEQUENCE:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) CONFIGURATION: linear
(xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO. 24

Arg Met Glu Arg Val Leu Arg Glu Asn

CLAIMS

1. Recombinant polypeptide comprising at least one active portion derived from a polypeptide, natural or synthetic, biologically active, genetically inserted in an albumin or one of its variants or derivatives.

2. Polypeptide according to claim 1, characterized in that the biologically active peptide has therapeutic activity and is of human origin.

3. Polypeptide according to claim 2, characterized in that the polypeptide having a therapeutic activity is chosen from among all or part of the enzymes, enzyme inhibitors, antigenes, antibodies, hormones, receptors, coagulation factors, interferons, cytokines, growth and/or differentiation factors, factors involved in the genesis/resorption of bone tissue, chemotactic factors, motility or cellular migration factors, cytostatic factors, bactericidal or antifungal factors, or plasmatic, interstitial or extracellular matrix adhesive molecules.

4. Polypeptide according to one of claims 1 through 3, characterized in that the polypeptide having a therapeutic activity is chosen from among any antagonist or agonist peptide sequence having molecular and/or cellular interaction involved in the pathologies of circulatory and interstitial compartments.

5. Polypeptide according to one of claims 1 through 4, characterized in that the active portion has a structure chosen from among:

a) the entire peptide structure or,

b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or more residues) and preserving therapeutic activity.

6. Polypeptide according to one of claims 1 through 5, characterized in that the active portion is inserted strictly into the interior of the albumin or surrounded by junction sequences.

7. Polypeptide according to one of claims 1 through 6, characterized in that the active portion is inserted preferably at the level of those regions on the albumin that presumedly form regions exposed to the surface of the molecule.

8. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 5 extending from residues 57 to 62 of the albumin.

9. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 8 extending from residues 103 to 120 of the albumin.

10. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 13 extending from residues 178 to 200 of the albumin.

11. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of the region from residue 415 to residue 425 delimited by the helices h2 and h3 of domain III in the albumin.

12. Polypeptide according to one of claims 1 through 12, characterized in that the active portion is inserted in a single or multiple manner.

13. Polypeptide according to claim 12 characterized in that the active portion is repeated several times in the same place and/or in different regions of the albumin.

14. Polypeptide according to one of claims 1 through 13, characterized in that the active portions inserted are different in nature.

15. Variant of a nucleotide sequence coding for the albumin or one of its variants or derivatives including at least one single non-natural restriction site.

16. Nucleotide sequence coding for a polypeptide according to any of claims 1 to 14.

17. Nucleotide sequence according to claim 16 characterized in that it comprises a "leader" sequence allowing the secretion of the expressed polypeptide.

18. Expression cassette comprising a nucleotide sequence according to claim 16 or 17 under the control of a transcription initiation region and possibly a transcription termination region.

19. Self-replicating plasmid comprising an expression cassette according to claim 18.

20. Eukaryote or prokaryote recombinant cell in which a nucleotide sequence according to claim 16 or 17 or an expression cassette according to claim 18 or a plasmid according to claim 19 has been inserted.

21. Recombinant cell according to claim 20 characterized in that it is a yeast, an animal cell, a fungus or a bacterium.

22. Recombinant cell according to claim 21 characterized in that it is a yeast.

23. Recombinant cell according to claim 22 characterized in that it is a yeast of the type Saccharomyces or Kluyveromyces.

24. Process for preparing a polypeptide as defined in one of claims 1 to 14 characterized in that a recombinant cell is cultivated according to one of claims 20 to 23 under expression conditions, and the polypeptide produced is recovered.

25. Pharmaceutical composition comprising one or more peptides according to any of claims 1 to 14.

26. Pharmaceutical composition comprising a nucleotide sequence according to either claim 16 or 17 that may be used in gene therapy.

CLAIMS

1. Recombinant polypeptide comprising at least one active portion derived from a polypeptide, natural or synthetic, biologically active, genetically inserted in an albumin or one of its variants or derivatives.

2. Polypeptide according to claim 1, characterized in that the biologically active peptide has therapeutic activity and is of human origin.

3. Polypeptide according to claim 2, characterized in that the polypeptide having a therapeutic activity is chosen from among all or part of the enzymes, enzyme inhibitors, antigenes, antibodies, hormones, receptors, coagulation factors, interferons, cytokines, growth and/or differentiation factors, factors involved in the genesis/resorption of bone tissue, chemotactic factors, motility or cellular migration factors, cytostatic factors, bactericidal or antifungal factors, or plasmatic, interstitial or extracellular matrix adhesive molecules.

4. Polypeptide according to one of claims 1 through 3, characterized in that the polypeptide having a therapeutic activity is chosen from among any antagonist or agonist peptide sequence having molecular and/or cellular interaction involved in the pathologies of circulatory and interstitial compartments.

5. Polypeptide according to one of claims 1 through 4, characterized in that the active portion has a structure chosen from among:

a) the entire peptide structure or,

b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or more residues) and preserving therapeutic activity.

6. Polypeptide according to one of claims 1 through 5, characterized in that the active portion is inserted strictly into the interior of the albumin or surrounded by junction sequences.

7. Polypeptide according to one of claims 1 through 6, characterized in that the active portion is inserted preferably at the level of those regions on the albumin that presumedly form regions exposed to the surface of the molecule.

8. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 5 extending from residues 57 to 62 of the albumin.

9. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 8 extending from residues 103 to 120 of the albumin.

10. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of region 13 extending from residues 178 to 200 of the albumin.

11. Polypeptide according to one of claims 1 through 7, characterized in that the active portion is inserted at the level of the region from residue 415 to residue 425 delimited by the helices h2 and h3 of domain III in the albumin.

12. Polypeptide according to one of claims 1 through 12, characterized in that the active portion is inserted in a single or multiple manner.

13. Polypeptide according to claim 12 characterized in that the active portion is repeated several times in the same place and/or in different regions of the albumin.

14. Polypeptide according to one of claims 1 through 13, characterized in that the active portions inserted are different in nature.

15. Variant of a nucleotide sequence coding for the albumin or one of its variants or derivatives including at least one single non-natural restriction site.

16. Nucleotide sequence coding for a polypeptide according to any of claims 1 to 14.

17. Nucleotide sequence according to claim 16 characterized in that it comprises a "leader" sequence allowing the secretion of the expressed polypeptide.

18. Expression cassette comprising a nucleotide sequence according to claim 16 or 17 under the control of a transcription initiation region and possibly a transcription termination region.

19. Self-replicating plasmid comprising an expression cassette according to claim 18.

20. Eukaryote or prokaryote recombinant cell in which a nucleotide sequence according to claim 16 or 17 or an expression cassette according to claim 18 or a plasmid according to claim 19 has been inserted.

21. Recombinant cell according to claim 20 characterized in that it is a yeast, an animal cell, a fungus or a bacterium.

22. Recombinant cell according to claim 21 characterized in that it is a yeast.

23. Recombinant cell according to claim 22 characterized in that it is a yeast of the type Saccharomyces or Kluyveromyces.

24. Process for preparing a polypeptide as defined in one of claims 1 to 14 characterized in that a recombinant cell is cultivated according to one of claims 20 to 23 under expression conditions, and the polypeptide produced is recovered.

25. Pharmaceutical composition comprising one or more peptides according to any of claims 1 to 14.

26. Pharmaceutical composition comprising a nucleotide sequence according to either claim 16 or 17 that may be used in gene therapy.

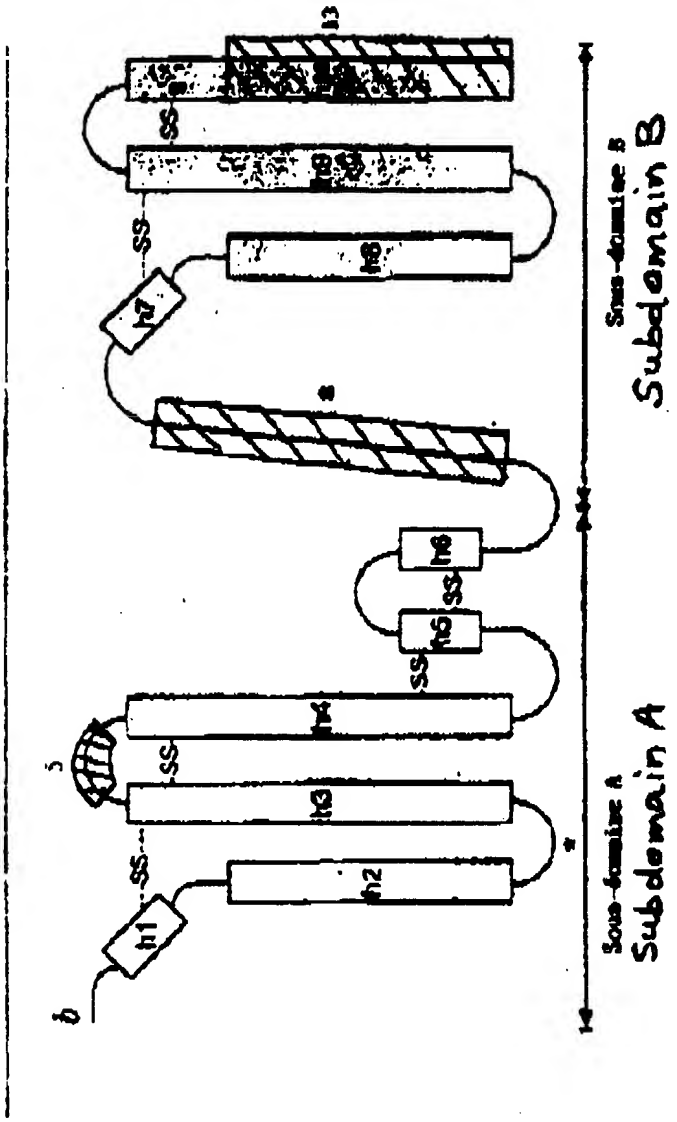
h3
SS
h8
h7
a
h6
SS
h5
SS
h4
b
SS
h3
SS
h2
h1
Sous-domaine B
Subdomain B
Sous-domaine A
Subdomain A

Figure 1

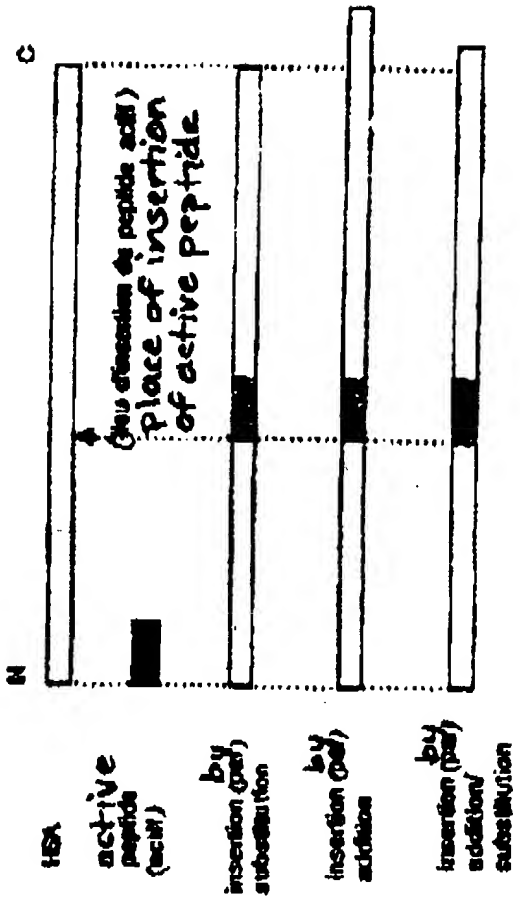
C
N
(site d'insertion du peptide actif)
Place of insertion of active peptide
HSA
active peptide (actif)
by insertion (pat) substitution
by insertion (pat) addition
by insertion (pat) addition/ substitution

Figure 2

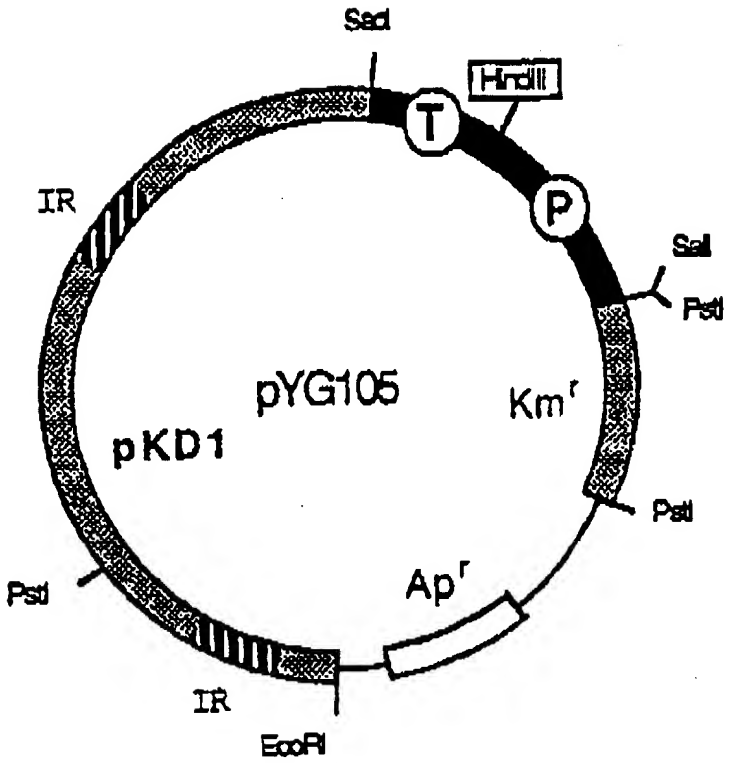
SacI
HindII
T
P
IR
SalI
PstI
pYG105
pKD1
Km^r
PstI
Ap^r
PstI
IR
EcoRI

Figure 3

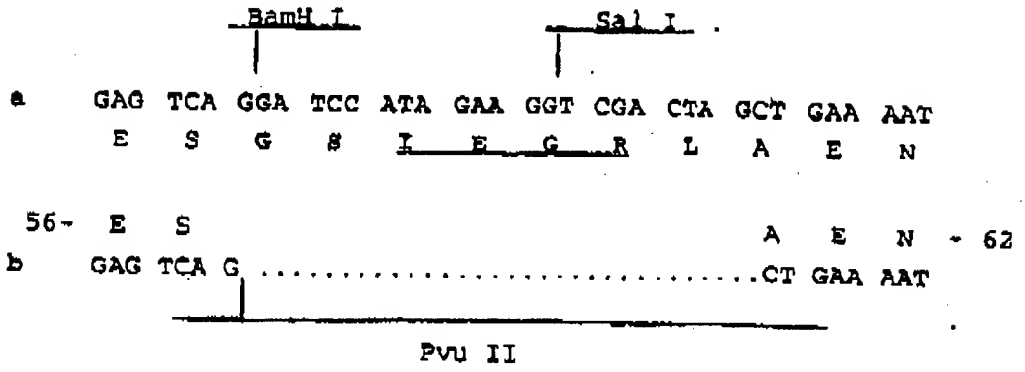
BamH I
Sal I
a GAG TCA GGA TCC ATA GAA GGT CGA CTA GCT GAA AAT
E S G S I E G R L A E N
56- E S A E N - 62
b GAG TCA G .....CT GAA AAT
Pvu II

Figure 4

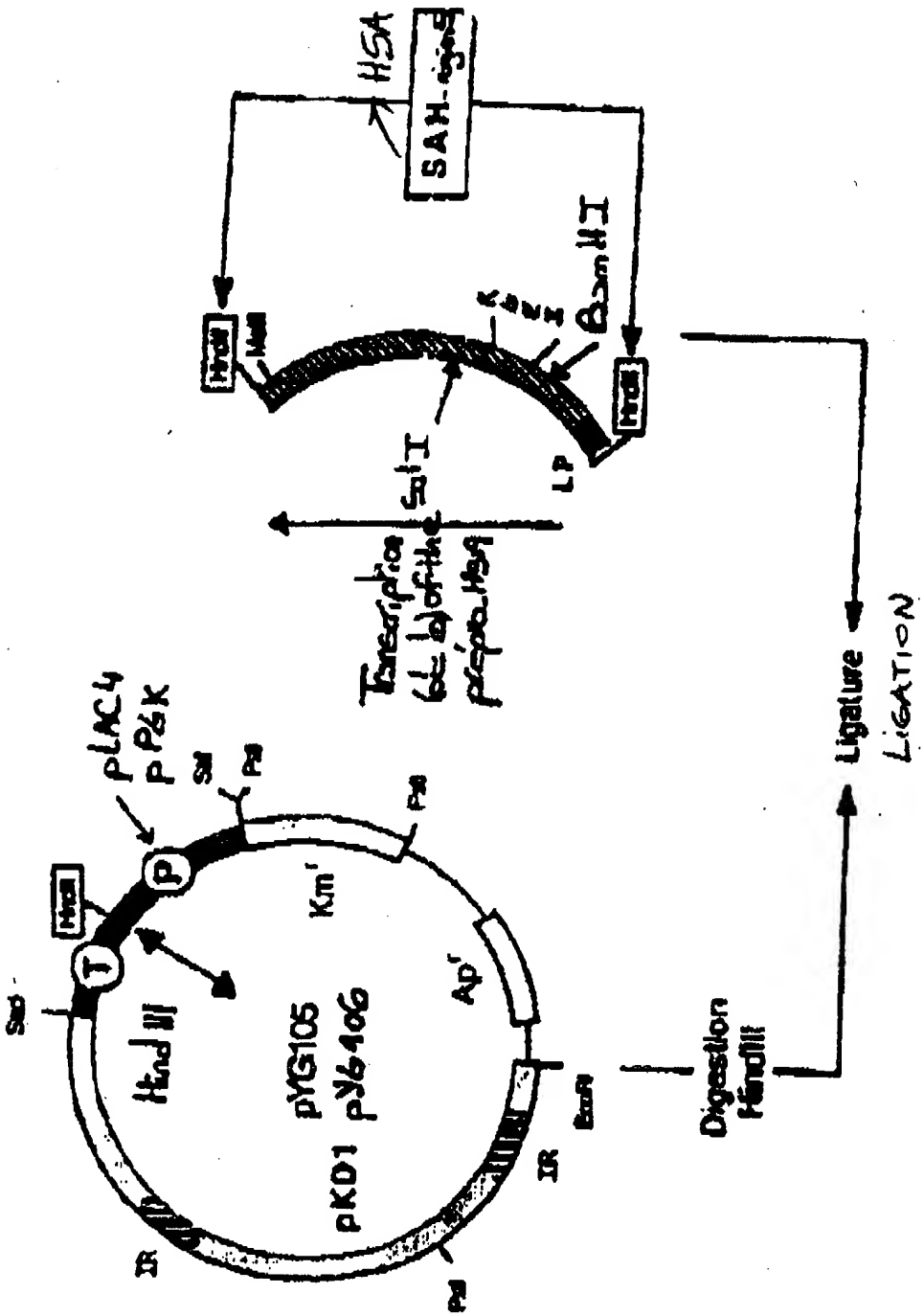
HSA
SAH-signal
BamHI
HindIII
LP
SalI
Transcription
pLAC4
PGK
Sal
Pst
Km'
Ap'
T
P
pYG105
pKD1
pYG406
HindIII
Sac
IR
IR
Digestion
HindIII
Ligature
LIGATION

Figure 5

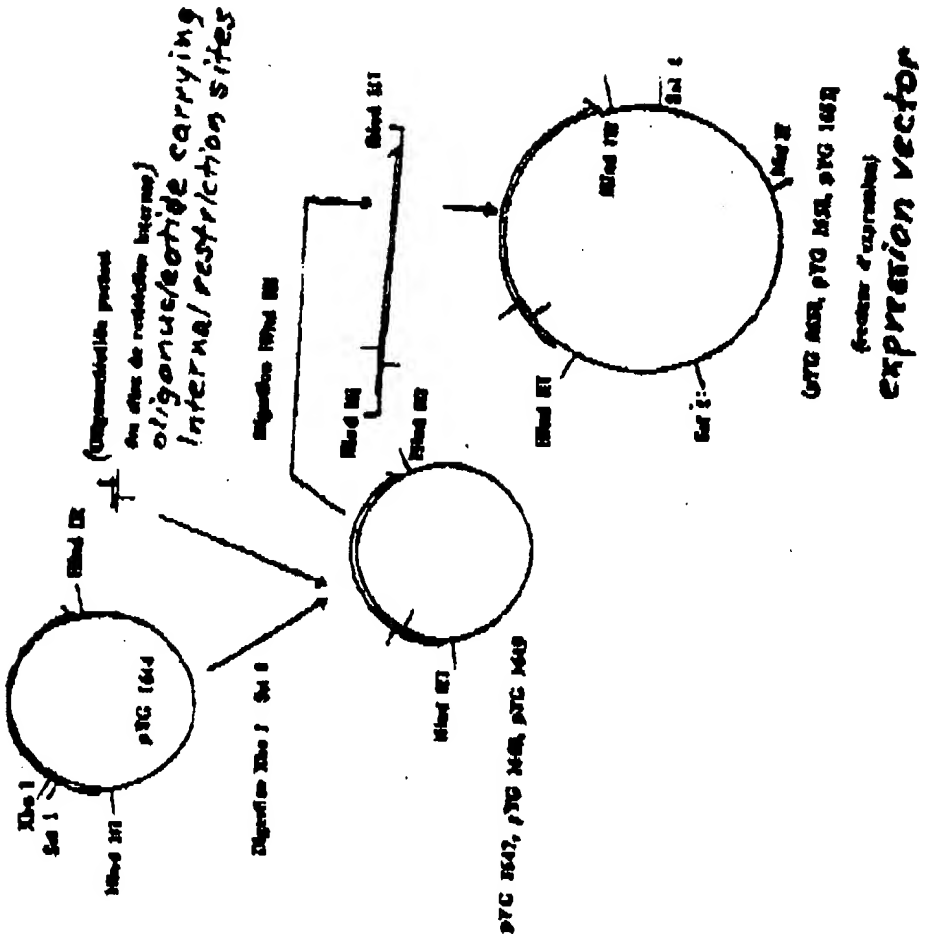
oligonucleotide carrying internal restriction sites
expression vector

Figure 6